DOCUMENT RESUME

ED 376 206 TM 022 318

AUTHOR Creemers, Bert P. M.; And Others
TITLE The International School Effectiveness Research Programme ISERP. First Results of the Quantitative Study.
PUB DATE Apr 94
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Case Studies; Comparative Analysis; Context Effect; *Cultural Differences; *Data Collection; *Educational Practices; Foreign Countries; *International Studies; Pilot Projects; Program Evaluation; Qualitative Research; *Research Reports; *School Effectiveness; Socioeconomic Status

ABSTRACT

The International School Effectiveness Research Programme (ISERP) is an example of the exchange of research and research results in the field of educational effectiveness. It aims to build on existing models of good practice and to avoid the variations in approach that limit the transferability of data within and between countries. A number of studies have been initiated as part of the overall program, including conferences, discussion papers, and the present pilot study that investigates outlier schools in different national cultural contexts. Countries involved are Australia, Canada, Hong Kong, Ireland, the Netherlands, Norway, Taiwan, the United Kingdom, and the United States. The qualitative part of this study includes case studies and a comparison of these studies among countries. The quantitative part of the study examines effective, averagely effective, and ineffective schools in each country, making a distinction between low socioeconomic status (SES) and middle SES schools. Data collection in 1992-93 and that planned for 1993-94 is yielding information that will elucidate variables within schools and contexts that affect school effectiveness. Five tables present preliminary data. (Contains 28 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN **OTHER THAN PAPER COPY** HAS BEEN GRANTED BY

DAVID REYNOLDS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED376206

THE INTERNATIONAL SCHOOL EFFECTIVENESS RESEARCH PROGRAMME

ISERP

FIRST RESULTS OF THE QUANTITATIVE STUDY

Bert P.M. Creemers
David Reynolds
Frans E. Swint

Paper presented at the Annual Meeting of the American Education Research Association
New Orleans
April 1994

TM022318

BEST COPY AVAILABLE

THE INTERNATIONAL SCHOOL EFFECTIVENESS RESEARCH PROGRAMME (ISERP): FIRST RESULTS OF THE QUANTITATIVE STUDY

Bert P.M. Creemers, David Reynolds, & Frans E. Swint

Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 1994

1 Introduction

The history of school effectiveness research can be described in different ways. A favourite way is to look at school effectiveness as a reaction to the quite pessimistic views on teachers, schools and education in general brought forward by the disappointing results of research. Another, quite different interpretation of the history of school effectiveness research considers this research as a natural prolongation of research from the past with respect to teaching, instruction, curriculum, school organization, and so on. Depending on one's views of history different godfathers for school effectiveness research are named, like Coleman et al. (1966) and Jencks et al. (1972). For most school effectiveness research the work of Edmonds (1979) and Brookover, Beady, Flood, and Schweitzer (1979) in the United States and of Rutter, Maughan, Mortimore, and Ouston (1979) in the United Kingdom are important starting points for school effectiveness research.

Especially in the United States, after Brookover and Edmonds a great deal of work has been done by researchers relating to the early American studies, but researchers in the United States seldom refer to any British or European literature other than that of Rutter et al. (1979) and Mortimore, Sammons, Stoll, Lewis, and Ecob (1988). In other countries researchers, although they refer to the earlier works of Brookover et al. (1979), Rutter et al., and Mortimore et al., show a similar ethnocentricity caused by the fact that in all these countries research was looking for factors that contribute to effectiveness within a specific country. The founding of the International Congress for School Effectiveness and Improvement and a journal associated with the organization (*School Effectiveness and School Improvement*) in 1988, stimulated the exchange of research and research results in the field of educational effectiveness. The International School Effectiveness Research Programme (ISERP) is one of the examples.

2 National and international school effectiveness studies

Out of the efforts of researchers in countries such as the United States, Great Britain, Canada, Norway, the Netherlands, Australia and New Zealand have come a large number of studies which look at the factors associated with student academic and social success

within different countries, and within specific communities or 'ecological niches' within countries (see reviews in Bashi & Sass, 1992; Creemers, Peters, & Reynolds, 1989; Reynolds, Creemers, & Peters, 1989). When we have a look at the national studies that have been conducted first, it is clear that these have:

1 Tended to focus almost exclusively upon only the cognitive or 'academic' outcomes of schooling, and to have neglected therefore the affective or 'social' outcomes of the schooling process.

2 Tended to be focussed more upon gathering 'process data' at the level of the school rather than upon the level of the classroom, where instruction is centred.

3 Tended to be atheoretical, and to have been concerned with only the establishment of relationships between variables rather than with the generation and testing of theories which would account for, and explain, those relationships.

4. Tended to ignore any possible variation in those factors associated with student learning within different cultural contexts within countries, preferring instead the use of 'steampress' varieties of 'whole sample' analysis which aggregate and look at relationships across all schools (the United States research from Louisiana would be a notable exception to this - see Stringfield & Teddlie, 1990; and Wimpelberg, Teddlie, & Stringfield, 1989).

5. Tended, until quite recently, to aggregate the data obtained on students, classrooms and schools, rather than explore any possible within student group, within class and within school variation (the development of varieties of multiple level methodology has made this aggregation now increasingly rare).

In addition to the defects in the national research bases that need to be remedied, the variation in the strength of the large number of national studies' research designs, research methodologies and in the mechanics of data collection in different societies makes the aggregation of the 'known to be valid' knowledge by which educational research progresses questionable, although recent effectiveness studies heavily depend on meta-analyses and reviews including international studies combine studies of different countries (Fuller & Clarke, 1994; Wang, Haertel, & Walberg, 1993). Within, and between, societies there are different conceptualisations of school and instructional process variables, different operationalisations even where the conceptualisation is similar and different measurement procedures and instruments even where the operationalisations are the same. Variations in sample characteristics, sampling techniques, methods of data collection, response rates, methods of analysis and statistical packages utilised increase the difficulty of the generation of any 'normal science' within the school effectiveness research community, and reduce the

possibility of any transfer of findings within and between countries at the level of educational policy. That much international educational policy discussion is currently simplistically based upon the transferability of factors shown to be effective in some cultures into the educational systems of others, merely increases the need to control the variation in the various national and international school effectiveness research enterprises. (The educational policy discussion may still be frighteningly simplistic even with more valid research findings, but that is not our problem!)

The findings of the nationally based school effectiveness studies themselves also create a pressing need for further research to explore the apparent inconsistencies and complexities within the existing bodies of knowledge. Assertive principal instructional leadership, for example, is one of the most replicated of school process factors within the American research literature in terms of being associated with positive student academic outcomes, yet within the Netherlands recent studies show no such principal effect (see review in Creemers, 1992).

Out of the efforts of international organisations like the IEA, and out of surveys by the Educational Testing Service entitled the International Assessment of Educational Progress (IAEP), have also come a wide variety of studies that have looked at the between country variation in student outcomes from schools, and at the factors that may explain variations between countries (IEA examples are Anderson, Ryan, & Shapiro, 1989; Postlethwaite & Wiley, 1992; Robitaille & Garden, 1989; Travers & Westbury, 1989. IAEP examples are Keys & Forman, 1989; Lapointe, Mead, & Phillips, 1989.).

The international studies of the IEA and the IAEP variety have virtually all of the defects of the national studies (cognitive bias in outcomes, and atheoreticism), together with some of the variability in the strength of research designs shown in the national studies, plus numerous defects of their own such as some of the greatest variations in response rates on items (between countries) that have ever been published. The measures of instructional and school processes tend to be rather simplistic, rudimentary and often to be organisationally and resource based. Variables at the level of the meso or macro societal contexts within which schools operate (the local and/or national economic structure, social structure and cultural structure, etc.) are not measured, so one has little understanding of the potential importance of these levels in potentiating or hindering the teacher and the schools' development of pupil academic and social development (Reynolds, Creemers, Bird, & Farrell, in press).

Then there is the issue of the 'black box' educational process areas into which neither 'nationally' nor 'internationally' based school effectiveness studies have been willing or able

to travel. The linkages between the levels of the classroom, or the instructional level, and the level of school processes is hardly explored, yet it is precisely the study of this interface which the use of, and findings from, multilevel methodologies has rendered necessary. Within-school variation by classroom in learning gains is now established to vary widely in some schools and less so in others - what 'interface' at the levels of the classroom and the school may be the explanation of this? We have at present no idea.

The educational factors outside the school at 'meso' or local community level are also relatively unexplored by the 'national' and 'international' studies except for the interesting work of Coleman and Laroque (1990). What factors within the district or local education authority or 'local educational state' have effects upon schools and their classrooms, in terms of potentiating or hindering pupil development? The effects of local economic, social and cultural meso context in terms of community influences, rather than the narrow educational context as above, also need investigation, particularly since some research shows considerable variation in what makes for 'effectiveness' in different ecological contexts (e.g. Wimpelberg et al., 1989). There are also whole groups of school process variables (e.g. school financial and physical resources) which have been historically neglected (in this case because of the now heavily dated literature reviews from the 1980s which used aggregated measures of resources and pupil outcomes, except for the recently published meta-analysis by Hedges, Laine, & Greenwald (1994), which suggested a larger influence of resources on educational outcomes than the famous analysis by Hanushek (1989)) and which need exploration.

3 The International School Effectiveness Research Programme (ISERP)

The ISERP programme of studies aims to do a number of things. It aims to build on existing models of 'good practice' in terms of research design, and aims to avoid the variation in national studies' research designs that limits transferability within and between countries, by utilisation of standard measures of intake, processes and outcomes, common methods of data analysis and common methods of data collection (although each participating country can add on to the common core any samples, instruments and analyses that may be appropriate for its own knowledge and/or policy needs).

It aims to improve on the quality of the national studies so far undertaken, and on the conceptualisation, operationalisation and measurement that is evident in the existing international studies. It aims to investigate those areas of school functioning where our ignorance has been revealed by the revolution of multilevel methodological techniques, and

where the changing national and international social structure may be generating new research and knowledge needs.

A number of studies have been initiated as part of the overall programme. Firstly, the past three years have seen an extended programme of meetings, the production of discussion papers and of presentations to many academic conferences, as we have sought to make sense of the school effectiveness literature, the international school/instructional effectiveness studies and the varying contexts, both socio-cultural and educational, of the different countries involved. Publications based on this phase have already appeared (Creemers, Reynolds, Schaffer, Stringfield, & Teddlie, 1991) and will shortly be appearing (Reynolds et al., in press).

Secondly, there is the present pilot study which attempts, through the study of outlier schools in different national cultural contexts, to both develop and test hypotheses concerning the variables at different levels (class, school, local context, societal context) which have an impact on student learning.

The third part of the ISERP programme is designed to increase our knowledge of effective instructional and school processes by linking together existing projects, including those of the IEA like the Third Mathematics and Science Study, that are ongoing naturally within different countries. This may involve trying to ensure that different projects share 'core' methodology and instrumentation, or alternatively may involve pooled analysis as appropriate. The aim is to make progress in the elucidation of those factors responsible for variation in education system performance.

4 A school outlier study within different countries

The countries involved in this study are Australia, Canada, Hong Kong, Ireland, the Netherlands, Norway, Taiwan, the United Kingdom, and the United States.

Certain characteristics, and principles, given our review and assessment of existing knowledge outlined above, can therefore be expected to be axiomatic in terms of the generation of our study design.

1 Data are being collected on individual pupils, on their classroom processes, school processes, on the classroom/school interface, and on the meso and macro factors that may potentiate or hinder student learning.

2 These data are on cohorts of students to whom their educational processes will be giving varying degrees of 'value added' educational and social development, with data being collected on students at intake and at outcome.

3 The outcomes of educational processes are seen as both academic and social in character.

4 Process variables are particularly intensive at the level of the instructional processes within classrooms where learning takes place.

5 Outside-school macro and meso level factors are being studied in detail, since they provide the educational and cultural contexts within which learning takes place.

6 Data collection is determined by the need to test models (and generate models) of school and classroom influences upon student learning and development.

In addition to those aspects of the study design that have been determined by our assessment of the principles which should guide us, more pragmatic considerations also determined that:

1 We should study younger rather than older children, since the older that children get, the greater is the chance of school effects being 'washed out' by being introduced into analyses as intake effects. Also, the younger that children are, the greater is the chance of them having no prior learning in the areas of knowledge that we are interested in.

2 We should maximise the variance on the dependent variables concerning student outcomes by choosing countries where there are apparently very high levels of pupil attainment (e.g. Taiwan) and also countries where there are apparently lower levels.

3 We should maximise the variance in outcomes within societies by choosing 'outlier' schools where there are high and low levels of 'value added' by the school (as well as choosing 'typical' or 'average' schools of course).

4 We should study educational effects upon curricular and knowledge areas where there is evidence of greatest school effects (e.g. science or mathematics), thus minimising the problems of controlling for home and outside-school influences.

5 We should choose as our focus a school subject (mathematics) that is the most uniform internationally in terms of knowledge content, to reduce the problems of measurement and comparability entailed with subjects where different societies teach different bodies of knowledge. However interesting it may be to examine the role of schools in the differential transmission of ideology through such areas as the humanities curriculum, research needs are for the definitions of subject content to be as near universal as possible across countries.

5 Research questions and design of the study

Generally speaking, the research questions can be phrased as follows:

1 Which factors are associated with student academic and social outcomes within a particular cultural context within each particular society within each country (e.g. within low SES context).
2 Which factors are in general associated with student academic and social outcomes across countries and what factors are restricted to a certain cultural context (e.g. within low SES context across countries, etc.).
3 Which factors are associated with academic and social outcomes across countries by different student characteristics.

More precisely, the research questions are as follows:

1 What are the differences between low effectiveness, middle effectiveness and high effectiveness schools in different countries?
2 What are the differences between schools in low SES and middle SES areas in different countries?
3 What factors contribute to effectiveness within countries?
4 What are the contributions of the different levels in the educational system to educational outcomes in different countries?
5 What are the differences between low effectiveness, middle effectiveness and high effectiveness schools between countries?
6 What are the differences between schools in low SES and middle SES areas between countries?
7 What factors contribute to differences in effectiveness between countries?
8 What are the contributions of the different levels in the educational system to differences in educational outcomes between countries?
9 Which are the contextual factors that can explain the differences between countries?

The study consists of a more quantitative and a more qualitative part. The qualitative part includes case studies within countries and a comparison of these case studies between countries. The quantitative study has a design that could be used for a larger longitudinal survey study, but is now applied to a small sample of schools. In each country effective, averagely effective, the so-called 'typical' schools, and ineffective schools are involved in this study. Within each group of schools a distinction is made between low SES and middle SES schools. From each of these categories at least one school, including two classes, is

involved. In some countries, like the Netherlands where schools normally do not have two classes per grade, just one class is included. Of course, this makes a comparison with respect to a distinction between class and school factors more difficult.

The selection of schools was based on an analysis of existing databases, when available. Otherwise, schools were selected on the basis of reputational criteria as identified by those key informers in the educational system who are likely to recognise positive, average and negative outliers. This can always be checked in the analysis of the data later on. The gainscore data will be used to validate the educational schools into the three effectiveness categories.

The data collection takes place at the individual student level with respect to student background, ability, and achievement in both academic and social areas; at the classroom level with respect to teacher behaviour, the curriculum and the organisation of the class; at the school level again with respect to the curriculum, the organisation of the school and the resources; and finally at the contextual level. The collection of data started in the schoolyear 1992/1993 and will be continued during the schoolyear 1993/1994, including research into the transition from one grade level to another. To make an international comparison not too difficult mathematics was chosen as an academic outcome measure. At the start the children were 7 years old, which implies that in most of the countries they are in the second grade of primary school where there is more emphasis on mathematics. Between countries, however, there are slight differences, such as in Norway where children start with primary education at the age of 7 and therefore do not have any mathematical background.

6 First results

The intake data were collected on mathematics attainment, non-verbal intelligence, occupation and educational qualifications of father/male caretaker, occupation and educational qualifications of mother/female caretaker and ethnicity. Student outcomes at the end of the schoolyear also included mathematics attainment in addition to other outcome measures, such as locus of control, attendance rate, language ability (a different test for each country), and attitude towards school, self, and mathematics. Table 1 presents reliability estimates for CTBS (Comprehensive Test of Basic Skills) and the IQ test taken in the first year. All coefficients were based on all students who took the test and for whom data at the item level were available. The given reliability coefficient is Cronbach's alpha, for the CTBS test this is the reliability based on the raw scores. Because these raw scores are transformed into scale scores, the estimates can also be used as indicators for scale score reliability. The CTBS consists of two parts: computations and application of mathematical concepts. The IQ test also consists of two parts: composition and exclusion.

In Table 2 the average scores for various countries are presented. The means are displayed and the standard deviations are given between brackets. The first CTBS score provides the intake results, the second one the outcome measure at the end of the schoolyear. The levels of CTBS indicate a higher difficulty level. The USA did not use the CTBS test. They used the CAT test instead. Therefore the raw scores on this test cannot be compared to the raw scores of other countries, but a transformation from CAT results into CTBS results is possible. The intake scores for Canada are based on one school. The outcome data for Taiwan and Australia are not available at the moment.

As expected, Hong Kong and Taiwan show high results in intake. The standard deviations are reasonable. The results of Great Britain, where mathematics education starts at the age of 5, are lower than in Hong Kong and Taiwan and the standard deviation is also slightly higher. The other countries took a lower level of the CTBS as an intake measure. This could be expected, especially for Norway, where, as already mentioned, school education starts at the age of 7. According to the CTBS, the raw scores at the various test levels can be transformed into one scale: the CTBS scale scores. These are presented in Table 3. Again, the means and standard deviation are displayed. The scale scores for the USA were computed by means of the transformation of the CAT scale scores into the CTBS scale scores. As expected, based on the raw scores, the scale scores for Taiwan and Hong Kong are the highest, although in the Hong Kong case the scale score for application in CTBS 13 is lower than in CTBS 12. The progression from one CTBS level to another probably causes problems. In addition to this, it can be questioned whether the CTBS is mirroring what happens in mathematics education in different countries. Countries may put emphasis on other aspects of mathematics education than measured by the CTBS. A further analysis of CTBS and mathematics education in the different countries is being carried out. In general, scale scores reflect what has already been indicated in the raw scores with respect to the means and the standard deviation.

Because of the fact that the transformation of raw scores into scale scores has caused problems, as the Hong Kong data show, we have refrained from computing gainscores on each test in order to calculate the learning effects for the grade level. First of all, a further analysis on CTBS and the transformation from raw scores into scale scores should be carried out.

To provide some insights into the effects of background variables, for each country the correlations between information on a student's background and CTBS scale scores were calculated. The results are given in Table 4. The information is provided per country, because computing the effects for all students might cause artefacts because of the different numbers of low SES students per country. The empty cells in Table 4 imply that no information was available on this issue. In the samples of Hong Kong and Taiwan there are

probably no children who can be described as ethnic minority. As can be seen in this table, there is a relationship between SES and the CTBS scores, but this is a rather weak relationship. The correlation is the lowest in Hong Kong, an indication of the strong support of school education by parents of all social groups? The 'medium' correlation between IQ and CBTS is an indication of the fact that IQ and mathematics are different things. As could be expected, there is a negative relationship between ethnicity and the CTBS intake.

The correlations between background variables and CTBS outcomes are mostly smaller. This is an indication that the influence of background variables is declining in favour of the influence of education in schools.

To carry out a quantitative analysis making use of the HLM multilevel analysis is problematic, because of the fact that in general the number of schools and classes in the countries are quite low and in some countries intolerably low. Nevertheless, to get an indication of how the variance between the classroom and the school level between and within countries could be split, such an analysis was carried out. But the interpretation has to be carried out with caution. However, the estimates can give a reasonable impression of the variance within each country's sample. As such the estimates can serve as a basis for the comparison of the various factors found in more qualitative material. For example, when there is no variance at the school level after correcting for intake in a given year, it is difficult to see why a school factor would have a substantial effect. These estimates, however, are less useful for within country generalisation of the magnitude of the influence of three levels. The main reason why they cannot be used for this purpose is that the schools were selected on an outlier basis, so as to maximise variance at the school level. This would mean a biased estimate of true school level variance in a country.

Table 5 presents the estimation of variance components for various countries for the computation part of the CTBS. An analysis was carried out on the CTBS raw scale scores. The first half of the table gives the estimates of variance with no predictors included, the second half of the table gives the estimates of the model with CTBS computation intake as predictor at the student level. In this way one can compare what part of the variance and outcome scores was also present at the intake. Table 5 presents the estimated variance and in brackets the number of units on which each estimate was based. As can be seen, especially at the school level the number of units was intolerably small. The initial estimates of the lack of variance for some countries might be due to the small number of schools. The adding of a predictor at the student level in this table explains variance at both the school and the class level. This is caused by the fact that the used predictor is also an achievement test (for outcome of the lower grade) which shows variance at the student, school and class levels.

It is interesting to see that in some countries the variance seems to cluster at the school level and in some countries at the class level. No class and school level predictors are included in

the model, and this can probably not be done in future, because of the small number of classes and schools in each sample. In the Netherlands most schools just have one class, so school and class are very much confounded with each other. The results seem to indicate that the class level is far more important than the school level, which goes in the same direction as the analysis carried out by Hill, Holmes-Smith, and Rowe (1993) and Scheerens, Nanninga, and Pelgrum (1989). Their research shows that the variance at the school level declines at the moment that the class level factors are introduced in a model that takes into account the intake. However, the number of units in our study is too small to take these findings as a further support of the results of earlier research.

However, we take this tentative analysis as an indication that we have to be careful with the factors at the school level and even more with respect to the contextual level.

Although the qualitative study can bring forth possible interesting variables between schools and context with respect to educational effectiveness, the contribution of these variables to the explanation of the dif' 'ences in student outcomes seems to be quite small. In order to find the amount of variance that would be interesting enough to be explained by these variables, a larger sample of schools, classes and perhaps also countries will be needed.

References

Anderson, L.W., Ryan, D.W.J., & Shapiro, B.J. (1989). *The IEA Classroom Environment Study*. Oxford: Pergamon Press.

Bashi, J., & Sass, Z. (Eds.). (1992). *School Effectiveness and Improvement: Proceedings of the Third International Congress for School Effectiveness, Jerusalem*. Jerusalem: The Magnes Press.

Brookover, W.B., Beady, C., Flood, P., & Schweitzer, J. (1979). *School systems and student achievement: schools make a difference*. New York: Praeger.

Coleman, J.S., Campbell, E.Q., Hobson, C.F., McPartland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

Coleman, P.J., & Laroque, L. (1990). *Struggling to be Good Enough*. Lewes: Falmer.

Creemers, B.P.M. (1992). School effectiveness and effective instruction: the need for a further relationship. In J. Bashi & Z. Sass (Eds.), *School Effectiveness and Improvement. Proceedings of the Third International Congress, Jerusalem* (pp. 105-132). Jerusalum: The Magnes Press.

Creemers, B.P.M., Peters, T., & Reynolds, D. (Eds.). (1989). *School Effectiveness and School Improvement. Proceedings of the Second International Congress, Rotterdam*. Lisse: Swets & Zeitlinger.

Creemers, B.P.M., Reynolds, D., Schaffer, E.C., Stringfield, S., & Teddlie, C. (1991). *International School Effects Research*. Kaohsiung: National Kaohsiung Normal University Press.

Edmonds, R.R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(1), 15-27.

Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Review of Educational Research*, *64*, 119-157.

Hanushek, E.A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, *18*(4), 45-65.

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes (An exchange: Part 1). *Educational Researcher*, *23*(3), 5-14.

Hill, P.W., Holmes-Smith, P., & Rowe, K.J. (1993). *School and Teacher Effectiveness in Victoria: Key Findings from Phase 1 of the Victorian Quality Schools Project*. Melbourne: Centre for Applied Educational Research, The University of Melbourne Institute of Education.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: a reassessment of the effects of family and schooling in America*. New York: Basic Books.

Keys, W.J., & Forman, D. (1989). *A World of Differences: A U.K. Perspective on an International Assessment of Mathematics and Science*. Oxford: Pergamon Press.

Lapointe, A.E., Mead, N., & Phillips, G. (1989). *A Word of Differences: An International Assessment of Mathematics and Science*. Princeton, N.J.: Educational Testing Service.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School Matters: The Junior Years*. Wells: Open Books.

Postlethwaite, T.N., & Wiley, D.E. (1992). *The IEA Study of Science II: Science Achievement in Twenty-Three Countries*. Oxford: Pergamon Press.

Reynolds, D., Creemers, B.P.M., Bird, J., & Farrell, S. (in press). School Effectiveness - The Need for an International Perspective. In D. Reynolds, B.P.M. Creemers, P.S. Nesselrodt, E.C. Schaffer, S. Stringfield, & C. Teddlie (Eds.), *Advances in School Effectiveness Research and Practice* (pp 183-201). Oxford: Pergamon Press.

Reynolds, D., Creemers, B.P.M., Nesselrodt, P.S., Schaffer, E.C., Stringfield, S., & Teddlie, C. (Eds.). (in press). *Advances in School Effectiveness Research and Practice*. Oxford: Pergamon Press.

Reynolds, D., Creemers, B., & Peters, T. (Eds.). (1989). *School Effectiveness and Improvement: Proceedings of the First International Congress, London, 1988.* Groningen: University of Groningen, RION.

Robitaille, D.F., & Garden, R.A. (1989). *The IEA Study of Mathematics II - Contexts and Outcomes of School Mathematics.* Oxford: Pergamon Press.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen Thousand Hours.* London: Open Books.

Scheerens, J., Nanninga, H.C.R., & Pelgrum, W.J. (1989). Generalizibility of instructional and school effectiveness indicators across nations; preliminary results of a secondary analysis of the IEA second mathematics study. In B.P.M. Creemers, T. Peters, & D. Reynolds (Eds.), *School Effectiveness and School Improvement. Proceedings of the Second International Congress, Rotterdam* (pp. 199-209). Lisse: Swets & Zeitlinger.

Stringfield, S., & Teddlie, C. (1990). School improvement efforts: qualitative and quantitative data from four naturally occuring experiments in Phases 3 & 4 of the Louisiana School Effectiveness Study. *School Effectiveness and School Improvement, 1,* 139-161.

Travers, K.J., & Westbury, I. (1989). *The IEA Study of Mathematics I - Analysis of Mathematics Curricula.* Oxford: Pergamon Press.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational research, 63,* 249-294.

Wimpelberg, R., Teddlie, C., & Stringfield, S. (1989). Sensitivity to context: the past and future of effective schools research. *Educational Administration Quarterly, 25,* 82-107.

Table 1 : Reliability of the instruments

Test Name	No of items	Reliability	No of students
CTBS 11 comp	28	0.92	562
CTBS 11 appl	34	0.88	562
CTBS 12 comp	30	0.91	2127
CTBS 12 appl	42	0.87	2127
CTBS 13 comp	38	0.84	449
CTBS 13 appl	42	0.76	449
IQ part 1	20	0.76	2526
IQ part 2	20	0.77	2526
IQ total score	40	0.83	2526

Table 2: Outcomes

	CTBS 11		CTBS 12		CTBS 13	
	comput	applic	comput	applic	comput	applic
Great Britain			15.5 (6.4)	23.5 (8.0)	17.3 (8.7)	20.2 (9.8)
Taiwan			25.9 (4.6)	32.9 (5.1)		
Norway	15.9 (6.4)	21.9 (6.5)	15.8 (5.9)	24.4 (8.0)		
Hong Kong			26.6 (3.5)	31.5 (5.3)	26.2 (5.9)	27.6 (6.8)
Canada	16.7 (6.2)	24.2 (7.6)	18.1 (5.1)	31.0 (5.9)		
Australia			17.8 (4.7)	28.7 (6.4)		
Netherlands	22.6 (4.7)	24.8 (4.1)	21.9 (5.4)	25.5 (6.5)		
USA						

Table 3: Scaled outcomes

	CTBS 11		CTBS 12		CTBS 13	
	comput	applic	comput	applic	comput	applic
Great Britain			504 (92)	572 (79)	578 (87)	612 (90)
Taiwan			629 (56)	657 (49)		
Norway	420 (109)	512 (92)	512 (80)	581 (74)		
Hong Kong			637 (42)	644 (48)	652 (41)	638 (43)
Canada	439 (83)	544 (116)	539 (64)	621 (70)		
Australia			540 (56)	619 (57)		
Netherlands	517 (71)	552 (53)	585 (61)	592 (57)		
USA	506 (85)	556 (101)	560 (62)	617 (73)		

Table 4: Correlations between variables

Country	Variable	intake comp	intake appl	outcome comp	outcome appl
USA	SES mother	.23	.26	.03	.00
	SES father	.25	.29	-.10	-.14
	Ethnicity	-.41	-.53	-.02	-.00
	IQ	.60	.59	.29	.42
Great Britain	IQ	.49	.58	.45	.41
Taiwan	IQ	.36	.54		
Norway	SES mother	.10	.02	.07	.18
	SES Father	.13	.05	.16	.07
	Ethnicity	-.30	-.30	-.34	-.29
	IQ	.38	.36	.44	.50
Hong Kong	SES mother	.05	.12	.14	.19
	SES father	-.07	.03	.07	.08
	IQ	.36	.54	.43	.45
Australia	IQ	.46	.51		
Netherlands	SES mother	.23	.27	.23	.15
	SES father	.11	.17	.05	.09
	Ethnicity	-.29	-.23	-.12	-.20
	IQ	.43	.57	.36	.49

Table 5: Estimation of variance components

	Empty Model			Model Intake predictor		
Country	Student Variance	Class Variance	School Variance	Student Variance	Class Variance	School Variance
USA	2437 (158)	460 (9)	0 (3)	1950 (158)	355 (9)	00 (3)
Great Britain	6243 (458)	197 (24)	1172 (12)	4532 (458)	0 (24)	423 (12)
Norway	5789 (399)	26 (19)	452 (9)	4590 (399)	0 (19)	163 (9)
Hong Kong	1490 (435)	210 (12)	0 (6)	1175 (435)	141 (12)	0 (6)
Nether-lands	2883 (134)	141 (8)	385 (6)	2003 (134)	130 (8)	0 (6)